AD P001341

Early identification of Academic Difficulties in Foreign Language Training

MAJ Allen N. Knox
Deputy Chief, Test Division
Defense Language Institute
Foreign Language Center

DLIFLC utilizes multiple sequential screening techniques for first-term enlistees of all Services who apply for foreign language training. Candidates must qualify on the AFQT, attain specified scores on selected parts of the ASVAB and attain a specified minimum score on the Defense Language Aptitude Battery. Despite these measures, academic attrition ranges from 12% to 30%, depending upon the language studied. A 56-item, five-option, attitude/interest instrument was constructed and administered to students before their courses started. Item leads centered on study habits, career goals, intercultural relationships and personal attributes. The five options varied from "strongly agree" to "strongly disagree." A scoring program permitted assignment of varying weights (4 to 0) to item options. For analysis, students were divided among three achievement groups based on 6th-week grades. This paper discusses the use of cross-tabulation tables (item option by achievement group) for item analysis and various option weighting strategies used to improve discrimination among achievement groups.

The purpose of this research was to develop a questionnaire which would complement the Defense Language Aptitude Battery in predicting those students of foreign language likely to encounter academic difficulty early in their course of instruction. The purpose of this paper is to describe some problems encountered in using cross-tabulation tables to weight item options to maximize predictive validity of the questionnaire.

Introduction: Enlisted students at the Defense Language Institute, Foreign Language Center (DLI) take at least three pencil-and-paper aptitude tests before being selected for attendance; Armed Forces Qualification Test, Armed Services Vocational Aptitude Battery (ASVAB), and Defense Language Aptitude Battery (DLAB). These tests and the nature of a volunteer military tend to cause the enlisted student body to be relatively homogeneous with respect to age and educational experience.

Despite selection procedures, academic attrition is approximately 14% and varies with the difficulty of the language being learned. Another predictive instrument (hereafter referred to as FLII) was deemed desirable to enable management to either (1) reduce attrition by applying different teaching strategies to those students whose DLAB and FLII scores indicated probable difficulty, or (2) to release early in the course those students whose DLAB, FLII and course scores indicated academic difficulty and probable attrition.

Management requirements for the FLII were that its development consume minimum resources, that it have a short administrative time (30 minutes or less) and that it be machine scoreable. Salient psychometric considerations were predictive validity and low correlation with DLAB.

Method: In order to meet the requirement of using minimum resources, it was decided to use items from past research projects. One source was the Foreign Language Interest Inventory developed at DLI in 1969. Because this instrument was designed to be administered to students during their course of instruction, many items were inappropriate for the new instrument or needed revision. A second source was a HumRRO study (Fiks and Brown, 1969) which had developed some items that correlated with academic success in learning foreign languages. The third source was a panel of experts who had a combined experience level in foreign language education and testing of 40 years. These three sources provided 56 items for the pilot form. Each five-option item described a factor (career goal, experience, intercultural relationships, personal attribute, etc.) thought to be related to learning a foreign language. The options typically took the form, agree strongly...not sure...disagree strongly, and the order of presentation was reversed on some items.

Items from the early Foreign Language Interest Inventory and the HumRRO instrument were originally assigned option weights of 1 (for one option) and 0 (for the remaining options). Traditional item analysis provided insight into which should be assigned 1.

Assigning option weights in this manner seemed undesirable for items purportedly measuring a continuum. A great deal of information may be lost when one option is weighted as "correct" and all others "incorrect". Likewise, for statistical analysis, a great deal of variability may be lost using this weighting scheme. Therefor, a scoring program was written for the FLII which permitted assigning a weight of 0 to 4 to each item option.

The FLII was administered to 279 students before they started their course of instruction. These students were randomly divided into two groups; a derivation sample and a validation sample. Due to administrative attrition and incomplete answer sheets, N=130 and 124 for the derivation and validation samples, respectively.

Course grade at the end of six weeks was used as the criterion. Students were divided into three academic groups according to course grade: >90=Hi, 80-90=Med and <80=Lo. Table 1. summarizes information on sample size and academic group.

TOTAL SAMPLE N=254

DERIVATION N=130

ACADEMIC GROUP	HI	MED	LOW
GRADES	91-100	80-90	0-79
N=	53	50	27

VALIDATION N=124

ACADEMIC GROUP	HI	MED	LOW
GRADES	91-100	80-90	0-79
N=	53	40	31

TABLE 1. SUMMARY OF SAMPLE

Item analysis was accomplished using 3 x 5 contingency tables (academic group by item option) generated for each item by Subprogram Crosstabs from the Statistical Package for the Social Sciences (Nie, et al, 1975). Fig. 1. shows "ideal" column percentages for an imaginary item measuring a continuous variable using five options. In this ideal example, half of the high (Hi) academic group chose option A. Other options were chosen by fewer members of the high group. The students doing relatively poorly in the course (Lo) most often chose options representing the opposite end of the continuum. The middle achievers (Med) most often selected options representing the middle of the continuum.

ACADEMIC GROUP

OPTION	LOW	MED	HI
A	0	5	50
B	10	20	25
C	15	50	15
D	25	20	10
E	50	5	0

Fig. 1. Contingency Table Showing "Ideal" Results-Column Percentages Only

Guttman (1941) showed that, for a quantitative criterion, the criterion mean of the examinees who chose that option could be used as the option weight. However, it was decided to use Subprogram Crosstabs because it was readily available and the output was in a form useful to other research.

Results: An a priori weight table, with options scored 0, 1, 2, 3, 4 or 4, 3, 2, 1, 0 depending upon how the options were presented, was used so that a high FLII score would be indicative of high academic grades. Using this straightforward weighting scheme, the FLII score correlated (Pearson) .2966 with sixth-week grades. In light of Cronbach's (1970) assertion that "Correlations of interests with grades in related fields are generally below 0.30...", the magnitude of this correlation was encouraging.

Contingency tables for each item were analyzed. The analysis indicated that "Lo students" and "Hi students" displayed rather similar option selection patterns. Unlike the pattern in Fig. 1., differences between Hi and Lo groups were a matter of degree rather than direction. Fig. 2. shows the contingency table (column percentages only) for item 37. Whereas 84% of the Lo group and 78% of the Hi group picked options A and B, 3 of 4 "Hi students" chose B and only 1 of 2 "Lo students" picked B.

	LOW	MED	HI	ORIGINAL WEIGHT	REVISED WEIGHT
A	33	28	6	4	0
B	51	56	73	3	4
C	11	16	20	2	3
D	4	0	0	1	0
E	0	0	0	0	0

Fig. 2. Item 37 Contingency Table, Column Percentages Only. Original and Revised Weights Shown at Right.

A similar pattern of responses was found for many items, and option weights were revised following these rules:

1) A high weight (usually 4) was assigned to that option chosen most often by the Hi group.

2) Lower weights were assigned to options when more Hi than Lo selected the option.

3) Zero was assigned to those options when more Lo than Hi selected it.

Note that no items were discarded at this time. Fig. 2. shows original and revised weights for item 37 based on this weighting strategy. FLII scores now correlated (Pearson) .09 with sixth-week grades.

Using five weights was notably unsuccessful. To reduce complexi and because responses of Hi and Lo groups appeared similar to warran five weights, a simplified weighting strategy was used:

1) A weight of 2 was assigned to options where tne percentage of Hi selecting an option exceeded the percentage of Lo by 20 or more.

2) A weight of 1 was assigned to options where the percentage difference was 10-19.

3) Zero was assigned to all other options.

The correlation between FLII score (using the new weighting) and sixth-week grades was .13. It rose to .24 when the Med group was excluded from calculation.

Since the Med group appeared to adversely affect correlation and had been ignored in weighting strategies, it was decided to include them in the weighting strategy. Fig. 1., depicting an "ideal" response pattern, shows that a pattern exists, not only within columns, but across rows.

In order to capitalize on the row and column pattern a new weight strategy was used:

1) Options A and E were assigned 4 or 0 if they closely matched the "ideal" row and column pattern. They were assigned 3 or 1 if they matched either row or column pattern.

2) Options B and D were assigned 3 or 1 if they closely matched the ideal row and column patterns or 2 if they matched only one pattern.

3) Option C weight was assigned dependent upon the pattern and weights of other options.

With this weight strategy, 7 items were discarded because of lack of any pattern. FLII score correlated .15 with sixth-week grades.

Thus far, all weighting strategies were inferior to the original, a priori weight table. It was decided to revert to the original weights, but refine them by assigning zero weight to those options chosen by 10% or less of the Hi group (See Fig. 3.). Seven items were deleted because response patterns showed no distinction between Hi and Lo groups.

	LOW	MED	HI	ORIGINAL WEIGHT	REVISED WEIGHT
A	74	77	83	4	4
B	18	18	13	3	3
C	7	4	3	2	0
D	0	0	0	1	0
E	0	0	0	0	0

Fig. 3. Item 24 Contingency Table, Column Percentages Only. Original and Revised Weights Shown at Right.

FLII scores now correlated .39 with sixth-week grades and .15 with DLAB scores. The correlation with grades regressed to .22 when the weight table was applied to the validation sample.

Conclusion: Gage (1957) found that partially a priori scores for five-option items were at least as valid for prediction as an elaborate weighting method. This research is partially supportive of those findings. While some improvement in predictive validity was possible using contingency tables for item analysis and variable option weighting, the method was trial-and-error and more time consuming than expected. The Guttman method is more rigorous and writing a computer program for its use does not appear difficult. DLI is presently writing such a program, and the results will permit a comparison between Guttman's method and the contingency table method used in this research.

References:

Cronbach, Lee J. Essentials of Psychological Testing, 3d Ed. New York: Harper & Row, 1970

Fiks, Alfred I. and George H. Brown. Student Attitudes and Foreign Language Learning. The George Washington University Human Resources Research Office, March 1969.

Gage, N. L. Logical vs. empirical scoring keys: the case of the MTAI. Journal of Educational Psychology, 1957, 48, 213-216.

Guttman, L. Supplementary Study B, pp. 251-364, in P. Horst (Ed), The prediction of personnel adjustment. New York: Social Science Research Council, 1941

Nie, Norman H., C. Hadlai Hull, Jean G. Jenkins, Karen Steinbrenner and Dale H. Bent. Statistical Package for the Social Sciences, Second Ed., New York; McGraw-Hill Book Company, 1975